

**DARPA-BAA-15-18**  
**Communicating with Computers (CwC)**  
**Frequently Asked Questions**

As of March 5, 2015

- Q17: Could you explain what is expected for the optional Abstract in the Statement of Work, Cost, and Schedule section, specifically regarding the cost estimate for resources over the proposed timeline of the project? Is a cost estimate the same as a cost proposal?
- A17: No, a cost estimate is not the same as a cost proposal. In the abstract, a total cost estimate is requested – this cost estimate is not binding and may be refined in your cost proposal submitted with your full proposal. No detailed breakdown of costs is expected for the abstract.
- Q16: Is natural language processing a required component of a proposed solution? Could I replace this with sign language for example?
- A16: It is up to the proposer to decide their own focus/modalities but there would be some concern whether systems can read/understand sign language with the same proficiency currently achievable for text and speech.
- Q15: The BAA says that information about the state of the blocks is available as "perceptual maps" -- it is not clear to me what these are?
- A15: A point of clarification – this phrase does not appear in the BAA. It appears in this FAQ, in the answer to Question 8. Given that, “perceptual maps” may be understood as describing what is visible from a particular point of view (i.e., the system’s perspective or the humans).
- Q14: Is the deliverable an end-to-end demo, or publications that describe the innovations/advances in state of the art? Is there a predefined task (e.g. collaboratively build a tower 5 blocks high), or is it more free form (build something that will be described in natural language at the time of the demo )?
- A14: As described in the BAA, there are three use cases. It is not required that a proposal address all three use cases. We anticipate that end-to-end systems will be built for the use cases by one or more performers, independently or collaboratively. Publications are also encouraged.

- Q13: Does use case 3 for collaborative composition of a narrative allow for dialogic behavior that includes negotiation of what the narrative arc will be, resolving misunderstandings, and other meta-dialog? Is it preferable to stick to an approach where a story will emerge from the moves of each partner in the dialog, or would it be of interest to include meta-level dialog where the agent and human user discuss characters' motivations, clarifications of causality, and so on.
- A13: Meta-dialog should not be confused with, substituted for, or considered to be of equal interest with dialog. That being said, it is up to each proposer to put forth the strongest plan for their own work.
- Q12: What sort of interaction will there be between performers of tasks TA2-TA4 and the teams who do TA1 and TA5?
- A12: DARPA will facilitate collaboration wherever possible. The needs of TA2-TA4 performers will inform the development of the apparatus provided by a TA1 performer. The capabilities and limitations of TA2-TA4 performers will inform the evaluation developed by the TA5 performer.
- Q11: For the collaborative composition use case, can we constrain the world of the narrative to a particular cast of characters and domain of objects (i.e., a specific database) as a means of linking the complex ideas of a narrative and a given lexicon to the world? An example would be the Harry Potter world versus the Percy Jackson world.
- A11: While one can imagine the usefulness of constraining the cast of characters and domain of objects to a particular source (such as the Harry Potter books), that is a more constrained use case that is desired. That being said, it is up to each proposer to put forth the strongest plan for their own work.
- Q10: If a team chooses to address a single use case, would it be necessary for the team to demonstrate that the technology is transferable to the other two use cases?
- A10: That would be desirable but not necessary. One of the premises of the program is that language is grounded in perception of the physical world. Given that premise, it is desirable that connections be shown between Blocks World and the other two use cases.

- Q9: Can non-US companies or university be part of a consortium or apply as a stand-alone entity?
- A9: This BAA is open to non-US participants. Please see Section III.A.2 of the BAA: "Non-U.S. organizations and/or individuals may participate to the extent that such participants comply with any necessary nondisclosure agreements, security regulations, export control laws, and other governing statutes applicable under the circumstances." In addition, please be aware that non-US prime contractors may, depending on the nature of the contract, be required to comply with US DOD accounting requirements beyond those they may be used to. For specific questions about these requirements, please write to CwC@darpa.mil.

As of February 26, 2015

- Q8: Can any additional information be provided regarding the apparatus that will be made available at the start of the Communicating with Computers program?
- A8: The apparatus provided by DARPA for the Blocks World use case will consist of a reasonably sized table instrumented with one or more Kinect sensors which will be used to track 15-20 uniquely identifiable cuboid blocks as well as a single human interacting with the "machine". The table will also house a TV screen or monitor at one end on which a simple cartoonish embodiment of the "machine" will be rendered. This will define the location and viewpoint of the "machine" to allow for communication where the semantics is colored by a perceiver's viewpoint. The screen may also be used for displaying any communication from the machine to the human, e.g., speech acts (textual) or non-verbal communication (gaze, facial expressions, etc.)

The system will provide a number of APIs, some of which have been described in the BAA, for accessing reasonably accurate information about the state of the blocks world as well as the human. The state of the blocks will be updated and reported as snapshots in time when the individual blocks' states each reach a stable state, i.e., after all motions or manipulations of the blocks have ceased. This will be reported both in terms of absolute coordinates and orientations of the blocks (the omniscient point of view) as well as perceptual maps from the viewpoints of the human and the "machine".

The communication from the human will be sensed and reported in synchronous streams for continuous signals like body pose, head pose and gaze directions and as asynchronous events for discrete signals like recognized gestures, and speech acts. A keyboard interface will also be provided to allow for directly inputting text based communication directed towards the machine.

The physical manipulation of the blocks from the machine's side can either be done via a human proxy who is expected to be someone other than the human involved in the interaction with the machine or using a robotic arm (not provided). Block manipulations

will be described in the same snapshot like language used to describe the Blocks World, i.e., define the goal state of a particular block resulting from the manipulation. Since such a representation will not be ideal for a human proxy to enact accurately, it is expected that such manipulation acts defined by the "machine" will be translated and presented visually to the human proxy.

Additional APIs, not mentioned in the BAA, will also be provided with the apparatus. First, a playback and review API will be provided that will give access to recorded game sessions. This will allow performers to review their sessions in great detail and may prove useful to TA5 performers when performing evaluations. A second API provided will allow streaming access to all raw data captured by the system (video, depth, audio, etc.).

Q7: Must a TA1 proposer propose to address all three use case? Or will there be one TA1 performer per use case?

A7: Proposers to the TA1 area are welcome to propose to all three use cases but it is not necessary. There will only be one TA1 performer per use case.

Q6: What is the phasing of the use cases? Will there be all three use cases throughout the program?

A6: We can easily imagine all of the use cases continuing throughout the program. Each use case will start as soon as there are performers ready to begin work, and each will continue for as long as it is useful.

Q5: In proposing a project on CWC's TA2 and/or TA3, is it necessary to focus on all three use cases?

A5: There are multiple alternatives for working on use cases, from focusing on one, two or three of them. It is up to the proposer to decide how to focus their efforts. As noted in the BAA, proposers are reminded not to overreach.

Q4: If DARPA will not release resources from maturing Big Mechanism technologies for awardees in the Communicating with Computers Program, is there any alternative to including a Big Mechanism participant in proposals for Communicating with Computers?

A4: There are multiple alternatives from teaming with a Big Mechanism performer to choosing not to propose work on the Biocuration use case. It is up to the proposer to decide how to address the issue.

Q3: If DARPA releases resources from maturing Big Mechanism technologies for awardees in the Communicating with Computers Program, what resources will be released?

A3: DARPA would facilitate access to Big Mechanism models and the ideas (hypotheses; questions; clarifications; etc.) produced by one or more Big Mechanism system regarding

those models. The CwC performer would be responsible for creating a communication channel from the human biocurator to the Big Mechanism system. Proposers might find it helpful to look at BioPAX/Pathway Commons models, which are one of the kinds of models being developed in Big Mechanism.

Q2: Will DARPA release resources from maturing Big Mechanism technologies for awardees in the Communicating with Computers Program?

A2: As needed.

Q1: Is the CwC program interested in the neurobiology of human-machine communication? Would it be within the scope of the program to understand the neurologic mechanisms that motivate people to interact with machines?

A1: The program has a more cognitive than neurological focus. TA2 and TA3 are about the construction and transmission of ideas. TA4 is about realizing the communication and includes functions that have been called discourse, dialog planning, pragmatics...That said, the question of what drives communication is certainly interesting and important. Keep in mind, though, that the CwC program seeks technology, so neurological accounts of aspects of communication that cannot readily translate to technology are less valuable than those that can.